This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):



BLACK BORDERS

- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.



11 Publication number:

0 079 739

A2

12

EUROPEAN PATENT APPLICATION

(21) Application number: 82305926.6

(22) Date of filing: 08.11.82

(51) Int. Cl.³: C 12 N 15/00

C 12 N 1/00, C 12 P 21/02 C 07 H 21/04, C 07 C 103/52 //C12R1/19, C12R1/865

(30) Priority: 12.11.81 US 320632

Date of publication of application: 25.05.83 Bulletin 83/21

(B) Designated Contracting States: BE CH DE FR GB IT LI NL SE (1) Applicant: THE UPJOHN COMPANY 301 Henrietta Street Kalamazoo, Michigan 49001(US)

(72) Inventor: Dugaiczyk, Achilles
c/o The Upjohn Company 301 Henrietta Street
Kalamazoo Michigan 49001(US)

(74) Representative: Perry, Robert Edward et al, GILL JENNINGS & EVERY 53-64 Chancery Lane London WC2A 1HN(GB)

(S) Albumin-based nucleotides, their replication and use, and plasmids for use therein.

(5) The DNA sequence coding for human serum albumin has been isolated and inserted as two fragments into two novel plasmids which can be replicated in *E. coli*. These novel fragments can be joined to provide a unitary DNA sequence which then can be cloned into a suitable host, e.g. *E. coli*, for the expression of human serum albumin (which is used extensively in medical practice in treating shock conditions).

EP 0 079 739 A

UPJOHN

٠ :

GJE 70/2056/02

ALBUMIN-BASED NUCLEOTIDES, THEIR REPLICATION AND USE, AND PLASMIDS FOR USE THEREIN

This invention relates to nucleotides related to human serum albumin (HSA), their replication and use, and plasmids (and host substances) for use therein.

The gene for serum albumin is regulated in 5 development. On the other hand, serum albumin is synthesised in mammals by the adult liver, and its plateau in The embryonic liver and yolk sac, on the adulthood. other hand, produce predominantly a-fetoprotein, but the synthesis decreases drastically after birth. Recently, 10 Law et al determined the complete sequence of mouse α-fetoprotein mRNA, Nature 291 (1981) 201-205. structure revealed extensive homology to mammalian serum albumin, indicating that the two proteins are encoded in the same gene family. Similar conclusions have been 15 reached from studies on the α -fetoprotein genes of the rat and the mouse; see Jagodzinski et al, Proc. Natl. Acad. Sci. USA, 78 (1981) 3521-3525, and Gorin et al, J. Biol. Chem. 256 (1981) 1954-1959.

The complete nucleotide sequence of human serum 20 mRNA has been determined from recombinant cDNA clones and from a primer-extended cDNA synthesis on the mRNA comprises 2,078 nucleotides, template. The sequence starting upstream of a potential ribosome binding site in the 5'-untranslated region. It contains all the 25 translated codons and extends into the poly(A) at the 3'-terminus. Part of the translated sequence codes for a hydrophobic prepeptide met-lys-trp-val-thr-phe-ile-serleu-leu-phe-leu-phe-ser-ser-ala-tyr-ser, followed by a basic propeptide arg-gly-val-phe-arg-arg. These signal 30 peptides are absent from mature serum albumin and, so far, have not been identified in their nascent state in humans. A remaining 1,755 nucleotides of the translated mRNA sequence code for 585 amino acids which are in agreement, with few exceptions, with the published amino 35 acid data for human serum albumin. The mRNA sequence verifies and refines the repeating homology in the tripledomain structure of the serum albumin molecule.

DETAILED DESCRIPTION OF THE INVENTION-

Human serum albumin cDNA is cloned into the Pstl site of plasmid pBR322 by the oligo(dG)-oligo(dC) tailing technique. Plasmid DNA was isolated from 97 positive colonies which hybridized to the enriched albumin cDNA probe, and the recombinant plasmid pHA36 was found to contain the largest insert of an albumin cDNA sequence. Its restriction endonuclease map is shown in the drawing, together with a restriction map of the primer-extended plasmid clone pHA206. The latter was obtained in a second transformation experiment after initiating the cDNA synthesis from an internal primer. This primer was a 91 base pairs long DNA fragment, MspI(152)-TaqI(182/3), isolated from pHA36. The two plasmids, pHA36 and pHA206, share 0.15 kb of homologous DNA. Together, they encode the entire sequence for human serum albumin, starting with the CTT codon for leu -10 of the prepeptide and extending into the 3'-untranslated region of poly(A).

Sequence of the Albumin cDNA. The sequence was determined for the most part on both DNA strands to ensure accuracy. All of the restriction sites used to end-label DNA fragments were sequenced across by 20 labeling a neighboring restriction site. The entire nucleotide sequence of the serum albumin mRNA, as determined from the cloned DNA in pHA36, pHA206, and from the primer-extended cDNA at the 5'-terminus of the message, is shown in the following Table 1. The inferred amino acid sequence is also indicated. The mRNA length is 2,078 nucleo-25 tides, of which 38 represent the 5'-untranslated region, 54 identify a prepeptide of 18 amino acids, 18 identify a propeptide of 6 amino acids, 1,755 code for the known 585 amino acids of serum albumin, 189 make up the 3'-untranslated region and 24 are the poly(A) sequence. Nucleotides 5 to 15 (-34 to -24) in the 5'-untranslated region (Table 1) are complementary to a 3'-terminal region of eukaryotic 185 RNA [Azad, A.A. and Deacon, N.J. (1980) Nucl. Acids Res. 8, 4365-4376] and thus could represent a ribosome binding site:

(5')...T
$$T^{C}T$$
 C T T C T G T......albumin mRNA
(3')...G A G G A A G G C G U C C $m_{2}^{6}A$ $m_{2}^{6}A$185 RNA

The translated portion of the mRNA sequence codes for the signal peptide and the main body of the albumin polypeptide chain. The

4083

signal peptide is composed of a hydrophobic prepeptide of 18 amino acids and a basic propeptide of 6 amino acids (Table 1). Since prepeptides are removed from nascent secretory proteins (like albumin) in the endoplasmic reticulum, they are seen only in vitro in heterologous translation systems. As yet, they have not been found within cells [Judah, J.D. and Quinn, P.S. (1977) FEBS 11th Mtg., Copenhagen 50, 21-29; and Strauss, A.W., Donohue, A.M., Bennett, C.D., Rodkey, J.A. and Alberts, A.W. (1977) Proc. Natl. Acad. Sci. USA 74, 1358-1362]. This is the first report of the presence and the sequence of a prepeptide for human serum albumin. As it is with other secretory proteins, the conversion of proalbumin to albumin takes place in the Golgi vesicles, and the enzyme responsible for this cleavage is probably cathepsin B [Judah, J.D. and Quinn, P.S. (1978) Nature 271, 384-385]. This is also a first report on the sequence of the propeptide for normal human serum albumin.

At the 3'-end of the message, the putative polyadenylation signal sequence, AATAAA, is located 164 nucleotides downstream from the amino acid termination codon TAA and 16 nucleotides upstream from the beginning of the poly(A) sequence. Another characteristic sequence located near the polyadenylation site has been identified by Benoist, et al. [Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids Res. 8, 127-142]; the concensus sequence from several mRNAs was concluded as TTTTCACTGC. A similar sequence, TTTTCTCTGT, is located 19 nucleotides upstream from the AATAAA hexanucleotide in the human albumin mRNA (Table 1).

TABLE 1

	(30)	(170)	(1560)	(350)	(440)	(330)	(629)	(710)	(300)
5	Ser AGC	20 1ys AAA	50 818 GCA	80 1eu CTT	110 Pro CCA	140 try TAT	170 Gln CAA	200 cys TCT	alu GAA
•	ohe TTT	phe TTC	phe TTT	thr	asn AAC	15th	169 cys TGC		230 lys ala glu phe ala glu AAA GCT GAG TTT GCA GAA
	18 CTC	asn	S X	ala GCA	ase GAC	tyr TAC	168 cys TCT	leu CTC	phe TTT
	-10 leu leu phe leu CTY CTY TTY CTC	16 84	thr	val	asp	thr phe leu lys lys ACA TTT TTG AAA AAA	ole GAA	ACA	phe pro lys ala glu phe TTT CCC AAA GCT GAG TTT
	C 18 C	ale GA	val GTA	75 gly asp lys leu cys thr val CGA GAC AAA TTA TGC ACA GTT	1ys AAA	lys lys Aaa aaa	ala ala phe thr GCT GCT TTT ACA	aln CAG	ala CCT
10	-10 lev lev CTT CTT	asp leu aly glu GAT TTG GGA GAA	e. A	£ 25 55	100 101 glu ays phe leu ain his lys GAA TGC TTC TTG CAA CAC AAA	leu TTG	ala phe thr GCT TTT ACA	1ys AA	1 ys
	ser TCC	leu TTG	val asn GTG AAT	75 1ya 1eu gya AAA TTA TGC	leu ain his TTG CAA CAC	phe	ala	ser ala TCT GCC	phe pro TTT CCC
	11e ATT		leu val asn TTA GTG AAT	1ys AAA	leu 11G	AC th	ala GCT		phe TT
	p r o trp val tlu phe TGG GTA ACC TTT	10 ale his erg phe lys GCT CAT CGC TTT AAA	leu TTA	asb CAC	phe TTC	glu qlu GAA GAG	1ys AAA	ala ser GCT TCG	gln arg CAG AGA
	t tu ACC	phe TTT	40 pro phe qlu asp his val lys CCA TTT GAA GAT CAT GTA AAA	aly ccs	101 0ys 7GC	a) n GAA	tyr TAT	ala GCT	aln CAG
15	val v	10 ala hís arg GCT CAT CGG	40 val GTA	70 thr leu phe ACC CTT TTT	100 asn glu AAT GAA	130 asp asn GAC AAT	160 ala lys arg CCT AAA AGG	190 asp qlu qly lys GAT GAA GGG AAG	220 leu ser CTG AGC
	trp TGG	his CAT	his CAT	Je CTT	AAT	ase GAC	17s	91.9 000	leu CTG
	-18 Met lys ATG AAG		glu asp GAA GAT	thr Acc	2 Z	1. L	als CCT	ala CA	919 CC
•		glu val GAG GTT	£ §	h is CAT	pro aly ara CCT GGG AGA	a LL	phe phe TTC TTT	asp	val ala GTA GCT
20	<u> CETTTICTCTTCTGTCAACCCCACAGCCCTTTGGCACA</u>		pro phe CCA TTT	CTT	glu pro aly ard GAA CCT GGG AGA	124 cys thr ala phe his rcc ACT GCT TTT CAT	phe TTC	glu leu arg GAA CTT CGG	lys ala trp ala val ala arg AAA GCA TGG GCA GTA GCT CGC
	755	lys ser AAG AGT	PT CCA	36T	gin glu CAA GAA	thr	pro glu leu leu CCG GAA CTC CTT	le CTT	ala GCA
	cccr	lys AG	34 gin gin cys CAG CAG TGT	1 ys			leu CTC	g lu GAA	lys ala trp AAA GCA TGG
	ACAG	ala hís GCA CAC	ole CAG	62 cys asp TGT GAC	ala lys GCA AAA	val met GTG ATG	ole AA	asp	1 23
	2000		gla			val GTG	970	leu CTC	
25	TCAA	-1 1 arg asp CGA GAT	Je CTT	60 glu asn GAA AAT	91 cys 1GT	asp GAT	150 tyr ala TAT GCC	1ys AG	210 ala phe GCT TTC
	retg	- £ 33	30 tyr		90 cys TGC	120 val GTT		177 180 ala cys leu leu pro lys GCC TGC CTG CCA AAG	210 ala CCT
	CTCT	p r o gly val phe arg GGT GTG TTT CGT	gln	ala GCT	asp GAC	ole CAG	phe TTT	leu TTG	gly glu arg a
	111	ag t	ala GCT	ser TCA	ala GCT	Pro CCA	tyr TAC	leu CTG	9 Y
	ક	val GTG	ala phe d	910	glu met GAA ATG	arg AGA	his pro CAT CCT	177 eys 160	91y CGA
30		91y 66T	နှင့် လ	asp	g lu GAA	val CTG	7. Y	a1 a	ã E
		arg AGG	11e ATT	ala GCT	aly GGT	leu 11G	arg AGA	asp lys ala GAT AAA GCT	gin lys CAA AAA
		-1 ser	150	val GTT	tyr TAT	87.9 CGA	arg AGA	lys AAA	5 ₹
		ser ala tyr TCG GCT TAT	leu val TTG GTG	53 thr cys val ACA TGT GTT	thr	5 7 200	al a		15 CT
35		ala GCT	leu TTG		g fu	leu CTC	11e	ata CCT	ser AGT
		367 700	21 ala GCC	51 173 AA	arg CCT	111 asn	# 3 8 8	171 ala CCT	201 ala GCC

	(890)	(980)	(1070)	(1160)	(1250)	(1340)	(1630)	(1520)	(1610)	(1700)
•	260 lev CTT	290 11e ATT	320 ala GrT	350 81a GCC	380 leu CCT	410 ard CGT	abo his CAT	&70 ser AGT	Syd 1 ys	530 val GTT
5	asp GAC	289 cys TGC	tyr TAT	leu CTT	pro CCT	val GTT	lys AAA	val GTA	ويو درد	led CTT
	arg ala AGG CCG	his CAC	asn AAC	ser vál val leu leu leu ara TCT GTC GTG CTG CTG AGA	lys AAA	ain asn ala leu leu val CAG AAT CCG CTG TTA GTT	438 cys TGT	5 Y	val GTT	هاه (۱۳
	arg	ser TCC	316 cys lys tgr AAA	leu CTG	phe TTT	leu CTG	437 cys TGT	thr	tyr val	thr
	asp GAC	lys AAA		leu CTG	asp glu GAT GAA	ele ACG	ser lys AGC AAA	lys AAA	glu thr GAA ACA	f &
10	ala asp CCT CAT	g g	asp val CAT GTT	val leu GTG CTG	asb GAT	AAT	AGC	ale GAG	₽. ¥3	1ys AA
		Jeu TTG		va] GTG	phe TTC	ala CAG	وا ک	5 5	asp GAT	1ys AAG
	253 glu cys GAA TGT	280 qlu lys pro leu leu glu lys GAA AAA CCT CTG TTG GAA AAA	ser lys AGT AAG	vál GTC	370 tyr ala lÿs val TAT GCC AAA GTG	lys phe AAA TTC	630 leu qly lys val CTA GCA AAA GTG	Jeu 17G	glu val GAA GTC	11e ATC
		pro CCT	ser AGT		1ys AAA	lys AAA	1ys AAA	val GTG		te de
•	250 Jeu leu CTG CTT	280 glu lys GAA AAA	glu GAA	340 asp tyr GAT TAC	ala SS	400 alu tyr GAG TAC	430 leu aly CTA GGA	460 461 leu cys TTA TGT	490 ala leu CCT CTG	ard AGA
15	250 Jeu CTG		310 val		370 tyr TAT	400 41c	tyo Lev CTA	460 1eu 11A	490 ala CCT	520 q1u GAG
	asp CAT	278 279 cys cys TGC TGT	asp phe CAT TTT	pro CCT	369 alu cys GAA TGC	9. 66.	AAC	o Po CAG	ser TCA	qlu lys GAG AAG
	5 59			h1s CAT		leu	879 AGA	AAC	phe TT	a) u
	P.I.S.	g S	ala GCT	arg AGG	7.55 CAT	ala CAG	ser TCA	CTG CTG	pro cys CCA TGC	ser TCT
20	345 246 cys cys TCC TCC	lys leu lys glu AAA CTG AAG GAA	ser leu ala TCA TTA GCT	ala arq GCA AGA	pro CCT	a la GAG	glu val GAG GTC	دها 130	pro CCA	518 cys thr leu TGC ACA CTT
20	345 glu cys GAA TGC	leu CTG	leu TTA		asp	phe TTT	g)u GAG	ser val	క్ట్ ప్ర	thr AGA
	9 8 8 8	AA A	ser	glu tyr GAA TAT	ala ala GCT GCA	le CTT	leu val CTT GTA	3e. 75	asn arg AAC AGG	514 11e cys ATA TGC
	thr	ser	pro CCT		ala GCT	o le	er c	CTA T		
	h is CAC	36. 100	asp leu GAC TTG	tyr TAT	ala OCC	392 cys TGT	pro thr CCA ACT	tyr TAT	leu val TTC CTC	ala asp GCA GAT
25	240 1ys val AAA GTC	270 ser 11e TCG ATC	asp GAC	1eu 776	361 cys TCT	asn	pro CCA	asp GAC	leu 11G	818 GCA
			300 ala GCT	330 phe TTT	360 cys 100	390 91n CAA	420 thr ACT	450 g1u GA	\$80 \$67	510 h1s CAT
	thr ACC	asp	pro CCT	met ATG	glu lys GAG AAG	lys AA	ser TCA	8 ts	9.€ 6&	phe TTC
	le CTT	gla	met ATG	917		11e ATC	val CTG	448 cys TGT	thr	thr ACC
	asp GAT	glu asn GAA AAT	glu	leu TTG	glu thr thr leu GAA ACC ACT CTA	asn leu AAT TTA	thr lys lys val pro gln ACC AAG AAA GTA CCC CAA	57 CC	677 673 TGC	phe TTC
30	ser lys leu val thr ICC AAG TTA GTG ACA		asp	phe TTC	thr	asn	979 CCC	met ATG	476 cys TCC	glu thr GAA ACA
	دي 160	265 cys 1GT	glu val glu asn GAA GTG GAA AAT	asp val GAT GTC	th ACC	gin CAG	val	87.9 ACA	arg val thr lys AGA GTC ACC AAA	£8 8
	leu 11A	lys tyr lle AAG TAT ATC	glu GA	asp	91u 6A	glu glu pro GAA GAG CCT	lys AA	lys AA	t hr	ala CCT
	1ys AAG	tyr TAT	val GTG	ala lys GCA AAG	tyr IAT	2 S	lys AAG	ار د کی	val CTC	phe asn TFT AAF
25		1ys AAG	91u GA	321 glu ala lys asp val GAG GCA AAG GAT GTC	351 1ys thr AAG ACA	9 to	th OC	441 pro glu ala lys arg CCT GAA GCA AAA AGA		
35	231 val GTT	261 ala CCC	291 ala GCC	321 914 GAG	351 1ys AAG	381 val GTG	411 tyr TAC	441 pro	471 asp CAC	501 91u GAG

	6	6	(2)	
	(1790)	(1883)) (20	
	SKO 173 AAG	₩Ċ	. ≹	
5	558 559 560 cys cys 1ys TCC TCF AAG	ter TTAA	10.77	
	558 cys 760	ACAT	AATT	
	1 ys	CATC	CATA	
	alu GAG	Ter TA	AAA(
	val	1 T A	CTAA	
10	550 pro lys ala thr lys glu gln leu lys ala val met asp asp phe ala ala phe val glu lys cys cys lys ccc AAG GCA ACA AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC GCT GCT TTT GTA GAG AAG TGC TGF AAG	567 570 cvs phe ala glu gly lys lys leu val ala ala ser gln ala ala leu gly leu ter tec cc cac cac cc tax axa axa ctt cct cca act caa cct tax cac cac cac cac ccac c	2161	
	ala GCT	TTA T	دلاوو	
	ala GCT	8 L8	CCAA	
	phe TTC	ala GCT	AAAG	
	550 asp CAT	580 91n CAA	CTCT	
15	asp CAT	AGT	GTTG	2078
	met ATG	818 GCA	1110	¥
	val	ala GCT	7.7.1	:
	ala GCT	val	TTT	20.
20	1ys AAA	Jeu CTT	ırcro	:
-	leu CTG	1 ys	TIC	₹
	9 Jn CAA	lys AA	XT1	MTC
	95	65.7 56.7	IAAAC	WC.
	lys AA	ale Ge	×10	755
25	560 Chr	570 91c SAG	IGAAC	(AAA)
	# 55 55	618 505	¥	MTA.
	lys AAG	phe TTT	X Y€	MTT.
	5 C	567 cys TCC	25	CTTC
20	lys AAG	thr	AATA.	TCTG
30	531 glu leu val lys his lys GAG CTC GTG AAA CAC AAG	asp asp lys glu thr GAC GAT AAG GAG ACC	ter ter cateragetaceateagaataagaaaatgaagateaaagettatteatetgitittettittegtiggtgtaaageceacetgtetaaaaacataaattetttaa (2002)	TCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAATGGAAAGAATCTAA 20AA (2078)
	1ys	lys AG	ACCA	CTTT
	val GTG	asp	.1333	CCCT
	leu CTC	asp GAC	CTCA	TTT
35	531 91u GAG	561 ala GCT	CAT	TCA

Following are examples which illustrate procedures, including the best mode, for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

5 Example 1 Isolation of Messenger RNA

Human liver mRNA was obtained following the procedure of Chirgwin, et al [Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) <u>Biochemistry</u> 18, 5294-5299]. Immunoprecipitation of albumin containing polysomes was performed according to Taylor and Tse [Taylor, J.M. and Tse, T.P.H. (1976) <u>J. Biol. Chem.</u> 251, 7461-7467]. <u>In vitro</u> translation of mRNA was carried out in a reticulocyte cell-free system, following the instruction of the manufacturer (New England Nuclear). The translation products were separated electrophoretically according to Laemmli [Laemmli, J.K. (1970) <u>Nature</u> 227, 680-685.

Example 2 Cloning Procedures

30

Double stranded cDNA was synthesized as described previously [Law, S., Tamaoki, T., Kreuzaler, F. and Dugaiczyk, A. (1980) Gene 10, 53-61]. It was annealed to PstI-linearized pBR322 DNA [Rolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crossa, J.H. and Falkow, S. (1977) Gene 2, 95-113] that had been tailed with 15 dG residues/3'-terminus [Dugaiczyk, A., Robberson, D.L. and Ullrich, A. (1980) Biochemistry 19, 5869-5873]. The annealed DNA was used to transform E. coli strain RR1, as detailed previously [Law, S., et al., Ibid.]. The albumin clones were selected using the colony hybridization method of Grunstein and Hogness [Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3965], with [32p]-labeled cDNA synthesized with the immunoprecipitated polysomal mRNA as template.

As shown in Example 5, plasmids pHA36 and pHA206 were deposited in <u>E. coli</u> HB101 hosts. The plasmids were obtained from <u>E. coli</u> RR1 hosts, described in this example, and transformed into <u>E. coli</u> HR101 by standard procedures well known to those of ordinary skill in this art. The <u>E. coli</u> RR1 hosts were lysed and then centrifuged to separate the chromosomal DNA, cell DNA and plasmid DNA. The plasmid DNA, remaining in the supernatant, is precipitated with ethanol and the precipitate is resuspended in buffer, e.g., TCM (10mM Tris·HCl, pH 8.0, 10 mM CaCl₂, 10 mM MgCl₂). The cells for transformation are

prepared as follows: 120 ml of L-broth (1% tryptone, 0.5% yeast extract, 0.5% NaCl) are inoculated with an 18 hour culture of HB101 NRRL B-11371 and grown to an optical density of 0.6 at 600 nm. Cells are washed in cold 100 mM NaCl and resuspended for 15 minutes in 20 ml chilled 50 mM CaCl₂. Bacteria are then concentrated to one-tenth of this volume in CaCl₂ and mixed 2:1 (v:v) with annealed plasmid DNA, prepared as described above. After chilling the cell-DNA mixture for 15 minutes, it is heat shocked at 42°C for 2 minutes, then allowed to equilibrate at room temperature for ten minutes before addition of L-broth 10 times the volume of the cell-DNA suspension. Transformed cells are incubated in broth at 37°C for one hour before inoculating selective media (L-agar plus 10 μg/ml tetracycline) with 200 μl/plate. Plates are incubated at 37°C for 48 hours to allow the growth of transformants.

15 Example 3 Mapping of Restriction Endonuclease Sites

Restriction endonucleases were obtained from Bethesda Research Laboratories and New England Biolabs and were used according to the manufacturers' instructions. The digested DNA fragments were analyzed electrophoretically on agarose [Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974) <u>J. Virol.</u> 14, 1235-1244] or acrylamide [Dingman, C., Fisher, M.P. and Kakefuda, T. (1972) <u>Biochemistry</u> 11, 1242-1250] gels.

Example 4 DNA Sequencing

DNA fragments were dephosphorylated with bacterial alkaline phosphatase (Worthington) and labeled at the 5'-ends with polynucleotide kinase (Boehringer-Mannheim) and γ[^{32p}]ATP. Following digestion with a second restriction endonuclease and electrophoretic separation of the fragments, DNA sequence determination was done according to the procedure of Maxam and Gilbert [Maxam, A. and Gilbert, W. (1980) Methods Enzym. 65, 499-560] and the degradation products were separated electrophoretically on 0.4 mm acrylamide gels as described by Sanger and Coulson [Sanger, F. and Coulson, R. (1978) FEBS Letters 87, 107-110].

Example 5 Recombinant Plasmids pHA36 and pHA206

As disclosed in Example 2, albumin clones were selected by hybridizing to the enriched albumin cDNA probe. Plasmid pHA36 contained the largest insert of an albumin cDNA sequence. Both plasmids pHA36 and pHA206 have been deposited in a viable E. coli host in the

permanent collection of the Northern Regional Research Laboratory (NRRL), U.S. Department of Agriculture, Peoria, Illinois, U.S.A. Their accession numbers in this repository are as follows:

HB101(pHA36) - NRRL B-12551

HB101(pHA206) - NRRL B-12550

5

<u>E. coli</u> HB101 is a known and widely available host microbe. Its NRRL accession number is NRRL B-11371.

NRRL B-12550 and NRRL B-12551 are available to the public. upon the grant of a patent. It should be understood that the availability of these deposits does not constitute a license to practice the subject invention in derogation of patent rights granted with the subject instrument by governmental action.

 $\underline{E.~coli}$ RR1 and $\underline{E.~coli}$ HB101 are known and widely available host microbes. Their NRRL accession numbers are NRRL B-12186 and NRRL B-11371, respectively.

pBR322 is a well known and widely available plasmid. It can be obtained from the following host deposit by standard procedures:

NRRL B-12014 - E. coli RR1 (pBR322).

YEp6 is a well known and widely available yeast episomal plasmid.

O It can be obtained from the following host deposit by standard procedures:

E. coli HB101 (YEp6) - NRRL B-12093.

Example 6 Assembly of the Serum Albumin Gene

Assembling the pieces together is a straighforward task of restriction enzymology. There is only one Mspl site in the overlapping
DNA sequence of the two cDNA clones. Two enzymatic steps of (i) Mspl
digestion of the two DNAs, followed by (ii) the use of ligase, an
enzyme that seals DNA fragments, will give the desired product.
Although two other undesired DNA species will also be obtained in the
course of this recombination reaction, both of them will differ substantially in size. Thus, separation and isolation of the desired DNA
species will be achieved.

The assembled DNA clone can be used to transform two types of cells:

(a) Escherichia coli

- (b) <u>Saccharomyces</u> cerevisiae
- (a) The vector of choice is plasmid pRR322, the same that has

been successfully used for cloning of the two fragmented pieces of the serum albumin cDNA.

(b) In order to transform yeast with the serum albumin structural gene sequence, the DNA must be inserted into one of the existing yeast plasmid vectors. This can be accomplished by taking advantage of the fact that several restriction endonuclease recognition sequences are absent from the cloned serum albumin DNA. Synthetic EcoRl DNA linkers can be ligated to the DNA fragment containing the serum albumin sequence followed by insertion (ligation) into one of the yeast plasmid vectors, e.g., YEp6, at the Eco Rl cloning site. The fused chimeric plasmid can be used to transform yeast according to an established procedure [Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. USA, 75, 1929]. YEp6 can be obtained from the NRRL repository, as disclosed supra.

15 Example 7 Expression of the Serum Albumin Gene

The main body of the structural gene will be transcribed by the E. coli or yeast enzymes. If little or no albumin is produced with the selected host, then an Escherichia coli promoter DNA sequence carrying an initiation codon, i.e., ATG, can be ligated at the begin-20 ning of the serum albumin structural gene. Such elements are known and available, e.g., lac promoter used for the expression of human interferon gene in E. coli [Proc. Natl. Acad. Sci. 77, 5230 (1980)]; source of promoter DNA [Proc. Natl. Acad. Sci. 76, 760 (1979)]. Also, see Nature, Vol. 281, October 18, 1979. It has already been 25 documented that such Escherichia coli promoter sequences function well in the expression of foreign genes in Escherichia coli [Mercereau-Puijalon, O., Royal, A., Cami, B., Garapin, A., Krust, A., Gannon, I. and Kourilsky, P. (1978) Nature 275, 505; and Goeddel, D.V., Kleid. D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Grea, R., Hirose, 30 T., Kraszewski, A., Itakura, K., and Riggs, A. (1979) Natl. Acad. Sci. USA 76, 106]. For expression in yeast, see Rose, M., Casadaban, M.J. and Botstein, D. (1981) Proc. Natl. Acad. Sci. USA 78, 2460 and 4466. Example 8 Screening of Clones Producing Albumin

Immunological methods can be used to detect small amounts of albumin made in a bacterium. Flat disks of flexible polyvinyl are coated with the IgG fraction from an immune serum and the disks are pressed onto an agar plate so that antigen released from an <u>in situ</u> lysed microbial colony can bind to the fixed antibody. The plastic

disk is then incubated with the same total IgG fraction labeled with radioactive iodine so that other determinants on the bound antigen can in turn bind the iodinated antibody. Radioactive areas on the disk expose X-ray film during autoradiography and thus identify colonies producing the protein which is being screened for. Detailed protocols of this procedure have been published [Broome, S. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA, 75, 2746]. The purification of human serum albumin can be accomplished by using procedures well known in the art. For example, procedures disclosed in a chapter by T. Peters: Purification and Properties of Serum Albumin, in: The Plasma Proteins, Putnam, Ed. Academic Press, New York, 1975, can be used.

The work described herein was all done in conformity with physical and biological containment requirements specified in the NIH Guidelines.

CLAIMS

- l. Plasmid pHA36, having a restriction endonuclease pattern as shown in the drawing.
- 5
 2. Plasmid pHA206, having a restriction endonuclease pattern as shown in the drawing.
- 3. E. coli HB101 (pHA36) having the deposit accession number 10 NRRL B-12551.
 - 4. $\underline{\text{E. coli}}$ HB101 (pHA206) having the deposit accession number NRRL B-12550.
- 5. A microorganism modified to contain a nucleotide sequence coding for the amino acid sequence of human serum albumin; said nucleotide sequence is as follows:

20

25

	(30)	(170)	(1560)	(350)	(440)	(330)	ししごソノ	(710)	(300)
	ACC	23 AA	8 8 8 GCA	8 2 E	110 Dro	140 177	170 aln caa	200 cys TGT	230 glu GAA (
5	ohe TTT	phe TTC	phe TT	thr	asn AAC	1 TA	169 Cys TGC	1ys AAG	ala GCA
ວ	<u>5</u> CTC	AAT	a Pr	818 CC	330	tyr TAC	168 cys 1CT	1ec CTC	phe TTT
	phe	g Jr GAA	thr		asb GAT	1ys AAA	₽ ₹	r s	9 e e
	C T E	₹ ₹	val GTA	thr val Aca GTT	1ys AAA	lys lys aaa aaa	thr ACA	aln ara CAG AGA	ala CCT
	ţ.	5 5	5.8		h is	phe leu TTT TTG	phe TTT	lys AA	1 ys
10		1e u 11G	asn	75 leu cys TTA TGC	al CAA		ala GCT	هاء درد	pro lys CCC AAA
		asp GAT	val GTG	1ys AAA	leu 11G	thr		ser TCT	ohe TTT
		lys AAA	15 T	88D GAC	ohe TTC	950	160 arg tyr lys ala AGG TAT AAA GCT	7.00 7.00 7.00 7.00 7.00 7.00 7.00 7.00	8 73 8 43
		phe	40 his val lys cat gta aaa	\$ 5 5	100 101 alu cys GAA TGC	130 asn alu AAT GAA	tyr TAT	190 lys ala AAG GCT	220 leu ser gln CTG AGC CAG
16		10 819 CCG	40 val GTA	70 leu phe CTT TTT	95 £ \$	130 asn AAT	160 arg AGG	190 173 AAG	220 ser gln AGC CAG
15		613 CAT	hIs CAT	CTT	AAT	asb GAC	1ys AAA	alu aly GAA GGG	arg leu CCC CTG
		ala GCT	85p CA1	th Acc	2 3	his CAT	a la CCT	£ \$	5 8
		val	₹.8	his CAT	g 1 y GGC	ala phe his GCT TTT CAT	ohe TTT	asp	ala GCT
		glu	1 phe	ser leu TCA CTT	pro	ala GCT	phe TTC	87.9 CCC	val CTA
20		AGT	pro CCA	ser TCA	g Ju	thr	1e CT 7	CTT	ele SS
		1ys AAG	34 gin cys CAG IGT	1 ys	£ &S	124 met cys ATG TGC	glu leu GAA CTC	918 GA	ala trp GCA TGG
		A S	919	asp GAC	1 ys	met ATG	eg. Sy	asp GAT	
		a1a GCA	gln	62 cys TGT	818 800	val	5 g	leu CTC	phe 1ys TTC AAA
		asp GAT	leu CTT	60 glu asn GAA AAT	91 cys TGT	asp GAT	a)a	1ys AAG	phe TTC
25		- F 50	30 tyr TAT		90 cys TGC	120 val GTT	150 tyr TAT	180 leu pro TTG CCA	210 ala GCT
		phe arg	gln CAG	ala CCT	SS CAC	glu	phe TT	1 eu	<u>د</u> کې
		phe TTT	ala CCT	ser TCA	ala CCT	5 Y	tyr	leu CTG	95
		val GTG	ala phe GCC TTT	g Ju GAG	met ATG		his pro CAT CCT	177 ala cys GCC TGC	
30		gly GGT	ala OCC	asp GAT	98	val	h1s CAT	1 818 CCC	ag I
30		arg ACG	leu ile TTG ATT	ala ser	91 <i>y</i> CGT	; arg leu val arg CGA TTG GTG AGA	ž ž	ala GCT	AA AA
		-1 ser TCC	1ed 77G	53 oys val rgr grr	tyr TAT		ala arg GCC AGA	1ys AAA	g) CAA
		-1 ser ala tyr ser rcg GCT TAT TCC	21 ala leu val leu (le GCC TTG GTG TTG ATT	53 0ys TGT	th	နို ည	- 818	asp GAT	ser leu gln lys phe gly AGT CTC CAA AAA TTT GGA
		ala CCT	leu TTG	51 lys thr AAA ACA	glu thr GAA ACC	Lec CTC	11e	ele CCT	ser AGT
35	,	106 207	21 818 GCC	21 173 AAA	81 819 CCT	111 888 AAC	161 glu 11e GAA ATT	171 ala ala esp GCT GCT GAT	201 ala GCC

	(890)	(980)	(1070)	(1160)	(1250)	(1360)	(1630)	(1520)	(1610)	510 phe his ala asp lie cys thr leu ser qlu lys glu arq qln lie lys lys qln thr ala leu val TTC CAT GCA GAT ATA TCC ACA CTT TCT GAG AAG GAG AGA CAA ATC AAG AAA CAA ACT GCA CTT GTT (1700)
5	260 Jeu CTT	290 11e ATT	320 ala GCT	350 818 GCC	380 1eu CCT	410 819 CGT	440 h1s CAT	870 Ser AGT	500 175 AAA (530 val GTT (
	asp GAC	289 cys TGC	tyr TAT	le. CTT	970 CCT	val GTT		val GTA	000	⊒ E
	260 arg ala asp leu AGG GCG GAC CTT	h is CAC	asn AAC	ara leu ala AGA CTT GCC	173	1 t	638 Cys lys TGT AAA	200	val p	2 2
	91.6 AGG	ser TCC		leu CTG	phe lys TTT AAA		637 Cys TGT			4 L
	35P (AC	1 ys		Jeu CTG	alu GAA	ala leu occ etc	lys o	1ys t	thr tyr ACA TAC	¥
10	asp asp GAT GAC	glu lys GAA AAA		340 asp tyr ser val val leu leu eu ara leu GAT TAC TCT GTC GTG CTG CTG AGA CTT	phe asp TTC GAT	asn a	430 pro thr leu val glu val ser arg asm leu qly lys val qly ser lys CCA ACT CTT GTA GAG GTC TCA AGA AAC CTA GGA AAA GTG GGC AGC AAA	qlu lys thr GAG AAA ACG		\$.₹
	8 T 8	1eu 11G	lys asp val AAG GAT GTT	va1	ale T	ala CAG	4 7 2	his car	val asp glu GTC GAT GAA	7 Y Y
	253 cys TGT	leu leu CTG TTG	AAG	val o	val GTG	phe o	val GTG C	leu his TTG CAT	al a TC G	5 F
	253 glu cys ala GAA TGT GCT	pro CCT	Ser	ser val TCT GTC	ala lys val phe GCC AAA GTG TTC	ξ. Υ.	lys v	val 1 GTG 7	glu val GAA GTC	- ₹
	leu CTT	1 ys	g tr	tyr TAC	ala CC 2	tyr 1ys TAC AAA	7 7 7	461 0ys v TGT 0	teu g CTG G	2 5 2 5
15	250 asp leu leu GAT CTG CTT	88 54 54 54 54	310 val	340 asp GAT	370 tyr :	aon glu tyr lys phe GAG TAC AAA TTC	430 leu qly CTA GGA	460 4 leu o TTA T	490 ala 1 GCT C	520 910 a GAG A
	asp	279 cys TGT	310 phe val TTT CTT	340 pro asp CCT GAT	369 cys TGC	91 7 7 2 2 2	ash J	91n 1 CAG T	ser a	43 S
	ફેસુ	278 glu cys CAA TCC		glu tyr ala arg arg his pro GAA TAT GCA AGA AGG CAT CCT	ale GAA	led CTT	879 A	460 461 ser val val leu asn gln leu oys TCC GTG GTC CTG AAC CAG TTA TGT	phe s	46 to
	2 to		pro ser leu ala ala asp CCT TCA TTA GCT GCT GAT	glu tyr ala arg arg his GAA TAT GCA AGA AGG CAT	h1s CAT	phe qlu qln leu TTT GAG CAG CTT	ser a	leu a	cys p	5 5
20	245 246 cys cys TGC TGC	leu lys CTG AAG	ala GCT	arg AGA	pro	glu GAG	val :	val 1 GTC 0	pro CCA 1	3 =
	245 246 val his thr glu cys cys GTC CAC ACG GAA TGC TGC	ser lys leu lys AGT AAA CTG AAG	leu ala TTA GCT	\$1\$ CCA	asp	phe TTT	glu val GAG GTC	ימן י כרכ (asn arg arg pi'o cys AAC AGG CGA CCA TGC	ر د ج
	548	ser lys AGT AAA	ser TCA	tyr	818 CCA		4a1	36T 7	arg arg AGG CGA	S14 cys t TCC A
	Acc		pro CCT	g P	ala SCT	glu leu GAG CTT	leu c	leu :	asn a	7 € 5
	1 1 1 2 CAC	ser TCC	1eu 11G		318 CCC	392 cys TGT	thr	tyr		1 de 1
25	240 lys val his AAA GTC CAC	270 ser 11e TCG ATC	300 ala asp cct cAc	Jeu 116	361 633 TGT	AAT	970 CCA		leu val TTG GTG	2 2
			300 a1a ccT	330 Phe TTT	8 2 8	390 gln CAA	420 thr	450 glv asp GAA GAC	480 ser 1 TCC 1	510 hts a
	asp leu thr GAT CTT ACC	asp	pro CCT	met ATG	glu lys c	173		818 GCA	glu s	1 k
	ct t	g y	met ATG	phe leu gly met TTC TTG GGC ATG	glu GAG	. .	val	648 0ys 1	thr 9	بر 5 م
30		glu esn GAA AAT	gle GAG	leu TTG	Iec CTA	gin asn leu il CAG AAT .TTA AT	gin val CAA GTG	פרט ב		i k
30	lys leu val thr Ang tta gtg aga	glu GAA	asp GAT	Phe TC	thr	AAT	. 679	met I	476 477 cys cys TCC TCC	ر د د
	val 616	265 cys TGT	asn AAT	val GTC	th So	gla CAG	val crA	5 2	lys c	\$ \$ \$
	lys leu AAG TTA	11e ATC	₹8	asp	38	pro CCT	lys val AAA GTA	lys arg AAA AGA	thr 1ys ACC AAA	5 F3
	Lys AG	tyr TAT	glu val glu asn asp GAA GTG GAA AAT GAT	lys AG	tyr TAT	976	lys AAG	818 CCA /	val t	Sn a
35	36. 700	261 ala lys GCC AAG		321 glu ala lys asp val phe leu GAG GCA AAG GAT GTC TTG	thr tyr ACA TAT	glv glu pro GAA GAG CCT	thr lys lys val pro gln val ser ACC AAG AAA GTA CCC CAA GTG TCA	648 glu sla lys arg met pro cys ala CAA GCA AAA AGA ATG CCC TGT GCA	arg val AGA GTC	he a
	231 val GTT	261 ala ccc	291 ala GCC	321 glu GAG	351 178 AAG	381 val CTG	411 tyr t	441 prog	471 839 8 CAC A	501 glu phe asn ala glu thr phe thr GAC TIT AAT GCT GAA ACA TIC ACC

	540 Ivs ala thr Ivs glu gln leu lys ala val met asp asp ohe ala ala phe val glu lys cys cys lys ANG GCA ACA AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC GCT GCT TTT GTA GAG AAG TGC TGC AAG (1790)	570 phe ala glu glu qly lys lys leu val ala ala ser gln ala ala leu qly leu ter TTT GCC GAG GGT AAA AAA CTT GTT GCT GCA AGT CAA GCT GCC TTA GQC TTA TAA CATCACATTTAAAAG (1883)	AAGAAATGAAGATCAAAAGCTTATTCATCTGTTTTTCTTTTTCGTTGGTGTAAAGCCAACACGTGTGAAAAAGATAAATTTCTTTAA (2002)
	0.50	ي	
	558 559 560 cys cys lys TCC TCC AAG	**	TT.
5	# 5 # # 5 #	ATTI	71.11
_	20 Z	ITCAC	TAAA
	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	ა გ	YYC
	53 55	7 te	AAAA
10	7 5 9 E	z Je	CTCT
10	有品	76 2 ⊀	700
	8 F	2 P	AACA
	9 E	18 81 17 60	YCC
	1 P	0 F V	CTAA
15	550 p + 55 IT CAT	580 r gln	16CT
	و <u>د</u> و پ	9 X	TCGT
	1 7 E	. a a l	1111
	3 5	2 S	71.
	₹8	2 E	1611
20	. 25 25	5 ₹	CATC
	2 5 2 2	ε ₹ ₹.	TAT
	20	7 Y B	AGCT
	50 ₹5	200	¥
25	S&O thr 1, ACA A	570 glu gl GAG GA	ş
	* * * *	ପ୍ରଧ	75
	* 9 * 9	- G	₹
			Š
	6 PT 0	567 r cys c 100	₽ ₹
30	• 1y	a G A G	ter ter TGAGAATAA(
	₹ 5 • • •	8 G	A ter
	531 glu leu val lys his lys GAG CTC GTG AAA CAC AAG	esp lys glu thr GAT AAG GAG ACC	ter ter CATCTCAGCCTACCATGAGAATAAGAGA
	* to	8 U	55
	- 20 C -	a asp	Ž,
35	5 g	561 ala CCT	3

TCATTTTGCCTCTTTTCTCTGTGCTTCAATTAATAAAATGGAAAGAATCTAA..... 20AA (2078)

6. Nucleotide sequence of the cDNA of human serum albumin, said nucleotide sequence is as follows:

			,					
5	(170)	50 ala GCA (260)	(350)	(440)	(330)	(620)	(710)	(300)
	20 1ys AAA	S	80 Jeu CTT	110 070 CCA	140 try TAT	545 545 545	200 cys TGT	230 glu GAA
	asn phe AAT TTC	alu phe CAA TTT	ala thr GCA ACT	AAC	leu TTA	169 cys 7GC	1ys AG	ala GCA
	asn AAT			330	tyr TAC	168 cys TGT	les CTC	a F
10	glu	val thr glu phe GTA ACT CAA TIT	thr val	asp CAT	Tys AA	5 AA	3.73 AGA	alu phe CAG TTT
	91u 6AA	val GTA		1ys AAA	1ys AAA	th ACA	a) a	ala CCT
	lys asp leu aly glu glu asn phe AAA GAT TTG GGA GAA GAA AAT TTC	g) u	75 88p lys leu cys GAC AAA TTA TGC	bhe leu ain his lys asb TIC TIG CAA CAC AAA GAT	leu TTG	phe TTT	lys AAA	1 ys
	Jeu TTG	AAT	leu TTA	a In CAA	phe	ala	313 500	pro CCC
	asp GAT	val GTG	1ys AAA	leu TTG	ACA A	ala GCT	ser ala TCT GCC	TT T
15			asp GAC		910	tyr lys Tat aaa	36T	ara AGA
	to arg phe CCG TTT	40 his val lys leu CAT GTA AAA TTA	70 phe qly TTT CGA	100 101 alu eys GAA TGC	a la	160 erg tyr lys ele ele AGG TAT AAA GCT GCT		a Ju CAG
	ာ ဥ			9 to 45	130 asn AAT	91 9 97 9 70 0	190 1ys AAG	220 Ser AGC
	ala his GCT CAT	asp his GAT CAT	thr leu ACC CTT	asn AAT	asp GAC	1 × ×	9 to 000	220 leu ser CTG AGC
20	ala GCT	asp GAT		ard AGA	ala phe his GCT TTT CAT	ala CCT	€ ₹	ala arg leu GCT CGC CTG
20	i asp ala his lys ser glu val GAT GCA CAC AAG AGT GAG GTT	o P	ser leu hís TCA CTT CAT	9 1 y GGG	phe TT	phe ahe TTC TTT	asp GAT	ala GCT
	976	pro phe CCA TTT	leu	ero CCT		phe TTC	819 000	val GTA
	ser AGT		62 oys asp lys ser TGT GAC AAA TCA	glu GAA	124 cys thr TGC ACT	leu CTT	glu leu GAA CTT	lys ala trp ala val AAA GCA TGG GCA GTA
	lys AAG	34 gln gln cys CAG CAG TGT	asp lys GAC AAA	g In CAA	124 cys 16C	leu		t. 166
25	hIs CAC	g In CAG	asp GAC	ala lys CCA AAA	val met GTG ATG	g tu	leu asp CTC GAT	al a
	ele GCA		62 9ys 1GT		val GTG	979 000		AA AA
	asp GAT	leu CTT	AAT	91 16T	120 glu val asp val met GAG GTT GAT GTG ATG	ှို့ ရှိ	180 Leu leu. pro 1ys CTG TTG CCA AAG	phe TTC
		30 tyr 1A1	60 glu GAA	90 cys TGC	120 val GTT	150 tyr 1A1	180 CCA	
		gla CAG	ala 601	asp GAC	ole GAG	phe TTT	1eu. 776	glu arg GAA AGA
30		als GCT	ser TCA	ala GCT	pro CCA	tyr TAC	leu CTG	£ 8
		9 t T T T	asp glu GAT GAG	glu met GAA ATG	g &	pro	17. 1975 1GC	9 l y 66 A
		ala phe GCC TTT	asp GAT	g Page	val GTG	h13 CAT	818 CCC	phe 111
		11e ATT	ala GCT	91 <i>y</i> cct	arg leu CGA TTG	arg AGA	ala GCT	1 ys
35		leu TTG	val	tyr TAT	879 CGA	8.5 A9.4	¥ 1	eg CAA
33		leu val leu lle TTG GTG TTG ATT	53 cys val TGT GTT	glu thr GAA ACC	5 TO COO	ala arg arg his pro GCC AGA AGA CAT CCT	asp	ser leu gln lys phe AGT CTC CAA AAA TTT
		Je u 776	th ACA	95 68	leu CTC	11e ATT	ala GCT	AGT
		21 ala GCC	27. 17.5 AA	B1 arg CCT	111 asn AAC	gle GA	171 ala CCT	201 ala GCC

5	(890)	(980)	(1070)	(1160)	(1250)	(1380)	(1630)	(0281)	(1610)	(1700)
	260 leu CTT	290 11e ATT	320 818 CCT	350 ala GCC	380 leu CCT	810 879 CGT	440 his CAT	870 8er AGT	500 175 AAA C	530 val CTT (
	arg ala asp AGG GCG CAC	289 cys 100	14	Je.	ere CCT	va1		val :	e o d	lea v
	\$1.8 800	h1s CAC	asn AAC	2 2	1ys AAA	7.7 T.	438 cys lys TGT AAA	0 to 0		ala leu GCA CTT
10		3e. 107	1ys AAA	Jeu CTG	phe TTT	Jeu CTG		thr a	tyr val	e ci
	asb asb GAT GAC	glu lys GAA AAA	316 cys TGC	leu leu ara leu CTG CTG AGA CTT	glu phe GAA TTT	8 18 CC	lys c			5 ¥
	asb GAT	pro leu leu glu 1ys CCT CTG TTG GAA AAA	val GTT	Jeu CTG	asb	asn a	430 ser arg asn leu qly lys val qly ser lys cys TCA AGA AAC CTA GGA AAA GTG GAC AGC AAA TGT	alu lys GAG AAA	alu thr GAA ACA	520 ser qiu lys giu arq qin ile lys lys qin thr ala TCT CAG AAG AGA CAA ATC AAG AAA CAA ACT GCA
	253 leú glu cys ala CTT CAA TGT CCT	Jeu 11G	asb GAT	val leu GTG CTG	phe TTC	ule CAG	9 y s		asb o	6 3 2 5
	253 cys TGT	Jec CTG		val GTC	781	9 e 1	val q GTG 0	leu h	val as	5 F
15	o S	973 CCT	ser lys AGT AAG	ser val TCT GTC	lys val AAA GTG	lys phe AAA TTC	1ys v	val l	glu val GAA GTC	= = = = = = = = = = = = = = = = = = =
	250 leu leu CTG CTT	280 glu lys pro leu leu GAA AAA CCT CTG TTG	alu GAA	tyr TAC	\$18		9,4	2 2 2	leu g CTG G	7 S
	250 leu CTG		310 val CTT	340 asp cat	370 tyr TAT	400 glu tyr GAG TAC	430 leu q CTA G	460 461 leu cys val leu his TTA IGT GTG TTG CAT	890 818 leu GCT CTG	520 glu 81 GAG AC
	asp GAT	279 cys TCT	phe TTT	970 CC1		4 y 29	asn 1		Ser a	6.9 X
20	91y GGA	278 cys TCC	asp CAT	CAT	369 alu cys GAA TGC	leu c	arg a	asn g AAC C	phe s	2 Z
20	h is	gle GAA	ala CCT	P SC	4 LS	aln J CAG (ser a	leu a CTG A	cys p	2 5 2 5
	245 246 cys cys TGC TGC	leu lys CTG AAG	ala	arņ AGA	670 CCT	olu GAG	val s	val 1 GTC C	pro c	3 E
		leu CTG	₹ 15	giu tyr ala arg arg GAA TAT GCA AGA AGG	asp GAT	phe o	pro thr leu val glu val CCA ACT CTT GTA GAG GTC	ser val val leu asn gln TCC GTG GTC CTG AAC CAG	arg pro cys phe CGA CCA TGC TTT	510 his ala esp ile cys thr leu ser qlu lys CAT GCA GAT ATA TGC ACA CTT TCT GAG AAG
	ale GA	ser lys AGT AAA	3er TCA	tyr TAT	313 GCA	leu r	val g	ser v	ະ ຄົນ	514 cys ti TCC A(
25	thr Ang		pro CCT	glu tyr GAA TAT	ala GCT	91° 646	leu v CTT (leu s CTA T	680 ser leu val asn arg TCC TTG GTG AAC AGG	× 0 ≥ ≤
	val his GTC CAC	lle ser ATC TCC	leu TTG	tyr TAT	818 CCC	392 cys 1GT (the J	tyr 1 TAT C	# Y 21 Y 2 Y 31	39 LI
			33	330 phe leu TTT TTG	361 cys TGT	AAT	pro t	asp t	2 C	# 3 2 %
	240 173 AAA	270 ser 700	300 313 CCT	330 phe TTT	360 cys TGC	390 91n CAA /	\$20 thr p	450 glu asp GAA GAC	\$80 ser 1	510 his ala cat gca
	thr Acc	88 CAT	pro			1 ys	ser t		glæ s GAA T	5 phe hi
30	led CTT	gla CAA	met ATG	gly met CCC ATG	glu lys GAG AAG	11e 1	val s	448 oys ale TGT GCA	thr g	# N
	asp CAT					•	• -		33.7	phe thr TTC ACC
	th ACA	glu asn GAA AAT	asp glu GAT GAG	phe TTC	thr	asn leu AAT ITA		15 t	476 477 cys cys TGC TGC	47
	val GTG	265 cys TGT	glu asn GAA AAT	val	glu thr thr leu GAA ACC ACT CTA	gln asn CAG AAT	val p	arg met pro AGA ATG CCC	∓0 ¥ ₹\$	glu thr GAA ACA
35	leu 77A	11e ATC	glu asn GAA AAT	rsp SAT	916		? \$	£ ₹ \$ ¥		1 G Q
33	1ys A6	tyr	val CTG	lys .	tyr g	glu pro GAG CCT	lys lys	ala lys CCA AAA	i AG	asn ala AAT GCT
	ser lys leu val thr asp TCC AAG TTA GTG ACA GAT	261 ala lys tyr lle GCC AAG TAT ATC	glu val GAA GTG	321 glu ala lys asp val phe leu GAG GCA AAG GAT GTC TTG	5r t	g ple g	411 tyr thr lys lys val pro gln TAC ACC AAG AAA GTA CCC CAA	441 pro glu ala lys arg met pro CCT GAA GCA AAA AGA ATG CCC	476 477 arg val thr lys cys cys AGA GTC ACC AAA TGC TGC	501 glu phe asn ala glu thr phe thr GAG TTT AAT GCT GAA ACA TTC ACC
	231 val CTT	261 813 600	291 ala c	321 glu a cAG G	351 1ys thr AAG ACA	381 val 9	411 tyr th TAC AG	441 pro gl CCT GA		c phe
				2.0	M - 4	m > U	# 2 F	ž	471 4SP CAC	501 91u CAG

5	558 559 560 cys cys lys TCC TCC AAG (1790)	580 ala ser gin ala ala leu qiy leu ter GCA AGT CAA GCT GCC TTA GCC TTA TAA CATCACATTTAAAAG (1883)	Ler ter ter Catctcagcctaccatgagaataaggaaaggatgaaggcttattcatctgtttttttt	
	560 1ys AAG	, ¥	TAA	
	558 559 cys cys TCC TCC	ter TTAA	TCT	
	558 678 700	ACAT	AATT	
10	1ys	CATC	CATA	
10	550 lys glu gin ieu lys ele val met asp asp bhe ele ele phe val glu lys c AAA GAG CAA CTG AAA GCT GTT ATG GAT GAT TTC GCT CCT TTT GTA GAG AAG 1	ter TAA	AAA	
	VB1	leu TTA	CTAA	
	2 E	580 phe ala glu glu qly lys lys leu val ala ala ser gln ala ala leu qly leu ter TTT GCC GAG GGT AAA AAA CTT GTT GCT GCA AGT CAA GCT GCC TTA GGC TTA TAA	CTCT	
	918 CCT	leu TT	נעננ	
15	8.58 8CT	200 818	V	
	phe TTC	ala GCT	LAAA	
	550 85p CAT	580 91n CAA	3GTG1	6
	asp GAT	Ser	CGTT	(207
	met ATG	2 to 000	Ë	₹.
20	va1 GTT	2 3 3 CCT	1101	
	sts GCT	val GTT	GT T T	20
	£.¥	leu CTT	ATCT	:
	leu CTG	1ys	ATTC	TAA.
25	gla	1ys	ECT.	¥¥
	914 GAG	qly GCT	¥¥	AAG
	₹ ₹	916	GATC	ATCG
	3 7 5	570 91u GAG	TGAA	AAA.
	lys ala	3. 200	WAAA	AATA
30		phe TTT	SAA40	W
	5 A	567 673 150	YGAG	CTTC
	lys AG	th.	te SATA	TCTC
	£ 5	910	ter	TCTC
35	531 giu leu val lys his lys pro GAG CTC GTG AAA CAC AAG CCC	567 ala asp asp lys glu thr cys GCT GAC GAT AAG GAG ACC TGC	YCCY.	TCATTITGCCTCTTTTCTCTGTGCTTCAATTAATAAAATGGAAAGAATCTAA 20AA (2078)
JJ	val 670	989 (CA1	(6661	ופככו
	l leu	,	rctc/	ATT.
	531 910 6AG	561 a1a GCT	3	5

7. Nucleotide sequence coding for the prepeptide of human serum albumin, said nucleotide sequence is as follows:

- · · · · · · · · · · · ·	3010	11001001106	sequence	1:
			(30)	
			ohe TTT	
			<u>5</u> 5	
			phe	
			ر <u>ہ</u> ہے۔ 12 کے 13	
			1 3 5	
			3er 700	
			116	
			1	
			7 45	
			2 % D	
			ys AG T	
			-18 -18 -16 1	
			¥ \	
			7565	
			נככדו	
			ACAG	
			ວວວວ	
			STCA	
			-18 p r o' Met lys trp val tlu phe lle ser leu pee leu phe ser TTTTCTCTTCTGTCAACCCCACACACCTTTGGCACA ATG AAG TGG GTA ACC TTT ATT TCC CTT CTT TTT CTC TTT AGC	7
			ודכדכ	•
			CCTT.	L
			_	•
				¥
				y- -1-
				•

er ale tyr ser arg gly val phe arg arg TCG GCT TAT TCC AGG GGT GTG TTT CGT CGA

8. Nucleotide sequence coding for pro human serum albumin, said nucleotide sequence is as follows:

5	(021)	(192)	(350)	(440)	(330)	(620)	(110)	230 ala qlu GCA GAA (300)
	20 1ys AAA	50 818 60A	80 1eu CTT (110 pro CCA (140 try TAT (170 aln CAA (200 cys 1GT (230 alu GAA (
	phe	phe 111	thr J	AAC (leu t	169 1 cys c TGC 0	1,38 AG 1	ele CCA
		ale CAA	818 GCA /	SSD S	tyr TAC 1	168 cys TGT	Jeu 1 CTC /	phe a
	alu XA	thr o		asb a	lys d	gle GAA	878 AGA 1	מוח מ
10		/al t	75 cas thr val	. × ×	2 X	thr q	oln a	عام دربا
	÷ 8	glu val GAA GTA	75 093 100 A	1 2 Y	leu lys TTG AAA	phe t	173 AA 0	Jys a
	leu aly alu TTG GGA GAA		3 ¥ 1 7	leu ain his lys TTG CAA CAC AAA			ser ala lys ain ara leu TCT GCC AAA CAG AGA CTC	pro 1 CCC A
	asp leu GAT TTG	val asn GTG AAT	lys leu AAA TTA	2 5 T	thr phe ACA TTT	ala ala GCT GCT	ser a	bhe pro TTT CCC
15			asp 1	phe 1	ale GAG A	8 X	ser s	
	10 arg phe 1ys CGG TTT AAA	40 val lys leu GTA AAA TTA	91y 8 CGA (101 6ys 9 1GC 1	glu glu thr phe GAA GAG ACA TTT	tyr lys TAT AAA	ala ser GCT TCG	220 leu ser gln arg CTG AGC CAG AGA
	10 879 000	40 val lys GTA AAA	70 phe g TTT C	64 45 6 44 6 44	130 asn g	160 arg t	190 173 8	220 3er 9 AGC C
	SAT O		leu r		asp a	- 8.₹	914	leu s CTG A
	10 ala his arg GCT CAT CGG	asp his GAT CAT	hr 1	arg asn AGA AAT	his a	25.0	alu aly GAA GGG	
20		glu asp GAA GAT	als (4) 500 CCC 1	phe h	3 5	asp q	ala arg GCT CGC
	glu val GAG GTT	phe c	leu his thr leu CTT CAT ACC CTT	pro CCT O	ale p	leu phé phe ala lys CTT TTC TTT GCT AAA	67.9	
	AGT	25	ser 1	glu pro GAA CCT	thr a	15	. Jeu a	ela val CCA GTA
	lys ser AAG AGT	34 cys pro TGT CCA	lys :	gin glu CAA GAA	124 cys (1GC /	leu 1 CTC (glu leu GAA CTT	£ 8
	CAC		asp lys ser GAC AAA TCA		met o	160 pro glu leu leu phe bhe ala lys arq CCG GAA CTC CTT TTC TTT GCT AAA AGG	asp g	ala trp CCA TGG
25	818 his GCA CAC	gin gin CAG CAG	62 oys asp TGT GAC	ala lys GCA AAA	val r	pro glu CCG GAA	leu a	lys AAA
	esp GAT	leu CTT	AAT		asp (1 ele	1ys	phe 1
	÷ 600	30 tyr TAT	8 gl 68	90 91 0ys cys TGC TGT	120 val GTT	150 tyr ala TAT GCC	180 pro CCA	210 ala p GCT
	r.9	977	ala	asp GAC	glu	phe TT	14c	25
30	o ag I	ala GCT	15 A	ala cct	ខ្ពស្ល	tyr p	10 TC	2 %
	p r o	phe	a n G	net :	leu val arg p	5.5	177 ala ala cys leu GCT GCC TGC CTG	gly glu arg
	919 507	ole con	asp CAT	glu met GAA ATG	/al a	arg his pro AGA CAT CCT	1	he g
	9-16 ACG	11e	ale CCT	gly cct	red 7	52	2 5	X 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
		716	/al	ryr (2 3	AGA A	lys AAA	¥ 1.
35		781 57G	53 cys val TCT CTT	thr tyr ACC TAT	pro arg	ala a	asp lys GAT AAA	leu gin lys phe CTC CAA AAA TTT
		leu ,	thr c	9 1.6 A	leu p	11e a ATT G	ela e ccr	ser 1 AGT C
		21 ala teu val teu Ite ata GCC TTG GTG TTG ATT GCC	173 t	81 3rg 9 CCT C	111 asn 1 AAC C	14 14 14 14 14 14 14 14 14 14 14 14 14 1	171 818 6CT 0	201 ala s GCC A

5	(890)	(980)	(1070)	(1160)	(1250)	(1340)	(1430)	(1520)	(1610)	530 asp lie cys thr leu ser glu lys glu arg gln lie lys lys gin thr ala leu val GAT ATA TGC ACA CTT TCT GAG ANG GAG AGA GAA ATC AAG AAA CAA ACT GCA CTT GTT (1700)
	260 1eu CTT	290 11e ATT	320 818 GCT	350 ala GCC	380 leu CCT (410 ara CGT (AAO his cat (2 - 0
	334	283 157 157	17.	35 T	pro 1	val a	1ys h AAA C		500 o lys c AAA	530 u val T CTT
	a1a 800	h is CAC	AAC .	£ 3	lys p	leu v TTA G	438 cys ly TGT A	pro val CCA GTA	1 pro T CCC	. e .
10	₽. 53 4		KA YS	5 E	phe 1ys TTT AAA	leu 1 CTG 7	437 & cys cy		r val	400
10	33		316 cys TGC /	5 5 1 5	glu p GAA T	ala 10	≆ 0 ¥ ≥ 5	s A A AG	thr tyr ACA TAC	A th
	CAT CAT	plu 1ys CAA AAA	316 val cys GTT TGC	eu 1	35 9 AT G	# Z	- E S	. o c A ₹	alu thr GAA ACA	5.3
	cys ala aso aso and ala TGT CCT CAT CAC AGG CCG	lec TTG	316 lys asp val cys lys asn tyr AAG GAT GTT TGC AAA AAC TAT	ser val val leu leu leu arg leu TCT GTC GTG CTG CTG AGA CTT	val phe asp glu GTG TTC CAT GAA	gin asn CAG AAT	¥ ¥	leu his aiu lys thr TTG CAT GAG AAA ACG	asp alu GAT. GAA	¥ 5
	253 cys TGT	leu leu CTG TTG	% % XG.	ים י זכ מ	# 5 T	phe a	2 5 2 5	⊋ 5		¥ * ×
15	_	פני	Ser AGT	ن به	2 X X	lys pl	₹ 5 2 5	val le GTG TT	glu val GAA GTC	7 7 1
	E Lea	1ys 1	916 CA A	tyr ser TAC TCT	18 J	F 5	23 25	48.5		2.9
	250 83p leu leu GAT CTG CTT	280 41u 6AA	310 val GTT 0	340 asp tyr GAT TAC	370 tyr ala lys val phe asp TAT GCC AAA GTG TTC GAT	400 glu tyr lys phe gln asn GAG TAC AAA TTC CAG AAT	437 ser arg asn leu qly lys val qly ser lys cys TCA AGA AAC CTA GGA AAA GTG GGC AGC AAA TGT	460 461 leu cys val leu hís qlu lys thr TTA TGT GTG TTG CAT GAG AAA ACG	490 ala leu GCT CTG	,
	SS CAT	279 ; cys TCT	phe r	pro a		9 7 50 50 4 9	2 - C			520 910 646
		278 0ys 1GC	ASD E	d s p	alu c	leu al CTT G	# ₹ 5 ≾	n gla c cAG	phe ser TTT TCA	¥ ;
20	245 246 cys cys his gly TGC TGC CAT GGA	278 Ber Ber lys leu lys glu bys TCC AGT AAA CTG AAG GAA TGC	ala a	ala arg arg his GCA AGA AGG CAT	369 818 818 829 pro hís glu cys GCT GCA GAT CCT CAT GAA TGC	gln leu CAG CTT	# ¥	tyr leu ser val val leu asn gln TAT CTA TCC GTG GTC CTG AAC CAG		28
	246 cys TGC	lys AG	ala e	5. Q	pro h		= 5 = 5	1 1eu	pro cys CCA TGC	<u> </u>
	245 246 glu cys cys GAA TGC TGC	55	lev pro ser leu ala TTG CCT TCA TTA GCT	ala arq GCA AGA	asp p GAT C	phe alu TTT CAC	glu val GAG GTC	1 val G GTC		F Le
	₹ 3	*X	ser leu TCA TTA	tyr a	ala a CCA G			r val C GTG	8 arg	\$ 5
25		150	25	glu tyr GAA TAT	ala a GCT C	glu leu GAG CTT	420 thr pro thr leu val ACT CCA ACT CTT GTA	tyr leu ser TAT CTA TCC	arg AGG	514 eys
25	240 leu thr lys val his thr CTT ACC AAA GTC CAC ACG	. S	300 ala asp leu pro GCT GAC TTG CCT	tyr g TAT G	ala a OCC O	392 cys gl TGT G	55		l asn	¥ ::
•	150	11e ATC	asp 1	leu t 776 7	361 cys a TGT 0		o thr A ACT	υ Στ	u val	25 Z
÷ . '	240 173	270 ser 1 TCG /	300 ala act	330 phe 1 TTT T	360 3 cys c; TCC TC	390 gin asn CAA AAT	T Pro	450 glu asp GAA GAC	980 ser leu val TCC TTG GTG	510 his ala cat cca
	A Fr	889 8	pro 8	3 met p	1ys c) AAG 10	173 91 AA C. 91			480 3er	510 h1s CAT
30	3 F	gla g	met p ATG C	91y m CCC A	glu 1) GAG AV		1 ser G TCA	s ala r ccA	. 9lu	phe
	asp .	ISN Q	glu met CAC ATC			. ∓ ₹ ≥ \$	n val A GTG	448 0 cys C 1CT	thr ACA	thr
	ser lys leu val thr asp TCC AAG TTA GTG ACA GAT	265 ays glu asn TGT GAA AAT	asp glu	phe leu TTC TTG	glu thr leu GAA ACC ACT CTA	asn leu ile AAT TTA ATC	pro gln ccc cAA	met pro ATG CCC	476 477 Cys cys TCC TCC	phe TTC
	18 7 7 7 7	265 0ys 9 1GT 0	37 A		. F O			met ATG	476 Cys TCC	45 \$
	7 € 7 €	1e o	# ₹ 2 \$	asp val GAT GTC	. A C	pro gln CCT CAG	6 val	giu ala iys arg GAA GCA AAA AGA	val thr lys GTC ACC AAA	glu GAA
35	lys leu AAG TTA	tyr ile TAT ATC	25	ىن ق ق	5 5	1 D	* 1 × 2 × 4 × 4 × 4 × 4 × 4 × 4 × 4 × 4 × 4	13s	thr	ala CCT
	ser l	45 t	2 S	~ * *	r tyr A TAT	glu glu GAA GAG	1ys	818 CCA	val GTC	asn AAT
	231 val s GTT T	261 ala lys GCC AAG	291 ala glu val glu asn GCC GAA GTG GAA AAT	321 glu ala lys GAG GCA AAG	351 1ys thr tyr glu thr AAG ACA TAT GAA ACC	381 val glu glu pro gln GTG GAA GAG CCT CAG	411 tyr thr lys lys val TAC ACC AAG AAA GTA	96 8	471 88p 8rg GAC AGA	501 glu phe asn ala glu thr phe CAG TTT AAT CCT CAA ACA TTC
	N 5 U	~ ონ	S & S	321 91u GAG	351 173 AAG	381 val CTG	411 tyr 1AC	841 pro CCT	471 38p CAC	501 91u CAG
					•					

5	559 560 cys lys TGC AAG (1790)	ter TTAAAAG (1883)	ITCTTTAA (2002)
10	558 : phe ala ala phe val glu lys cys or TTC GCT GCT TTT GTA GAG AAG TGC	leu qly leu ter TTA GGC TTA TAA CATCACATTTAAAAG (1883)	ter ter Catctcagcctaccatgagaataagaaaatgaagatcaaagcttattcatctgtttttcgttggtgtaaagccaacaccctgtctaaaaacataaatttctttaa (2002) Tcattttgcctcttttctctgtgcttcaattaataaaaatggaaagaatctaa 20aa (2078)
15	550 p asp phe ala ala T GAT TTC GCT GCT	580 r gin ala ala leu r CAA GCT GCC TTA	TGGTGTAAAGCCAACACC 178)
20	550 glu gin leu lys ala val met asp asp GAG CAA CTG AAA GCT GTT ATG GAT GAT	leu val ala ala se CTT GTT GCT GCA AG	20AA (20
25	LY8	570 phe ala glu glu qly lys lys leu val ala ala ser gln ala ala TTT GCC GAG GAT AAA AAA CTT GTT GCT GCA AGT CAA GCT GCC	AGATCAAAGCTTATTCA
30		567 thr cys phe ala gl	ter ter Catctcagcctaccatgagaataagagaaagaaatgaagatcaaagcttattcatctgttttttcgttgg TCATTTTGCCTCTTTCCTGTGCTTCAATTAAAAATGGAAAGAATCTAA 20AA (2078)
35	531 glu leu val lys his lys pro GAG CTC GTG AAA CAC AAG CCC	567 ala asp asp lys glu thr cys GCT GAC GAT AAG GAG ACC TGC	ter CATCTCAGCCTACCATGAC TCATTTGCCTCTTTTCTC

9. Nucleotide sequence coding for the pre pro human serum albumin, said nucleotide sequence is as follows:

5	(30)	(170)	(192)	(350)	(660)	(330)	(620)	(710)	(300)
	ACC	20 173 AAA	S = 5	80 leu CTT	110 670 CCA	180 try TAT	170 61n CAA	200 cys TCT (230 alu GAA (
	ag H	phe TTC	phe TTT	thr	asn AAC	leu TTA	169 cys TGC	lys AG	818 GCA
	3 CTC	AAT	alu CAA	ala CCA	35 CAC	tyr TAC	168 cys TGT	Je CTC	
10	phe TTT	glu	thr	val	aso CAT	lys AA	GAA	7 S	alu phe GAG TTT
••	2 E	atu GA	Val	thr		1ys AA	۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲. ۲	CAG	ala Grit
	12 5	el y SG A	£ 8	25 155	M S	leu lys TTG AAA	phe thr TTT ACA	₹ X	pro lys ala alu phe CCC AAA GCT GAG TTT
	3er 100	Je U	AAT	leu TTA		phe		318	919 000
	11e	SAT	val GTG	AAA	leu aln TTG CAA	ACA AT	818 GCT	101	phe 1
15	Phe TT	17. AA	leu val TTA GTG	asp lys	ohe TTC	5 g	1ys		ACA .
	r elu I tlu		40 val lys GTA AAA	ફેઇ	5 % 5 %	5 8	tyr lys ala ala TAT AAA GCT GCT	913 6CT	Ju s
	p r c trp val tlu TGG GTA ACC	10 arg phe CGC TTT		ope 111	충음중	130 88n AAT	160 AGC	190 lys ala ser AAG GCT TCG	220 ser gin arg ACC CAG AGA
		\$ 5 5	his CAT	le CTT	asn AAT	335 GAC			5 5
00	-18 Wet lys ATG AAG	#1# GCT	88D CAT	k K	2 5	his CAT	ala lys CCT AAA	alu aly GAA GGG	ala arg leu GCT CGC CTG
20	Het ATG	val GTT	₹ ₹	his CAT	, 55	phe TT	phe TTT	asp GAT	25
		g)u GAG	TT Phe	5 E	oro cct	818 CCT	phe TC	5 3	val GTA
		ser AGT	67 CCA	ser TCA	95 64	thr		CTT	2 S
		ly.	34 cys TGT	133	£ 442	124 673 160	Jec CTC	glu leu GAA CTT	trp 166
25		ala hís GCA CAC	914 CAG	62 asn cys asp lys ser leu his AAT TGT GAC AAA TCA CTT CAT	1ys	met ATG	glu leu leu GAA CTC CTT	asp GAT	ala SCA
			91,	62 cys asp 1GT GAC	# SC 43	va1 GTG	979 CCG	leu CTC	1ys AAA
		CAT CAT	Jeu CTT	asn AAT	92 253 761	asp GAT	8 1 8 OCC	1ys AG	T TC
		- £3	30 tyr	જ ટૂ જુ	90 91 cys cys 1GC 1GT	120 val GTT	150 tyr TAT	180 573	210 ala phe GCT TTC
		8.79 CGT	g] CAG	ala GCT	asp GAC	120 glu val GAG GTT	phe TTT	leu 116	5. A
30		9 E	ala GCT	ser TCA	ele CCT	i v	tyr TAC	1 e c CTC	3
		7 × 2 CTC	phe TTT	gle GG				177 eys TGC	÷ 5
		ply val	ala ccc	asp CAT	glu met GAA ATG	val CTG	his pro CAT CCT	177 ala cys GCC TGC	phe TTT
		9- 1- 90 V 00 V	11e ATT	st s	919	arg leu val arg CCA TTG GTG AGA	5 5	ala GCT	gin lys phe g CAA AAA TTT C
25		- , 5	125	val	tyr TAT	55	\$ \$	lys ala	-E X
35		als tyr ser GCT TAT TCC	val	53 cys val TGT GTT	th ACC	e D	ala arg arg OCC AGA AGA	esp GAT	
			160 11G	ACA th	918 8	lee CTC	11e	ala asp GCT GAT	Ser 1
		867 TCG	21 als GCC	S1 lys thr AAA ACA	914 CCT	111 837 AAC	# 26 X	171 818 607	201 ala ser GCC AGT

5	(890)	(980)	(1070)	(1160)	(1250)	(13%0)	(1430)	(1520)	(1610)	(1700)
	2&0 1eu CTT	290 11e ATT	320 ala CC:T	3Š0 818 GCC	380 164 CCT	410 8ra CGT	940 h13	470 ser AGT	Sno 1ys AAA	530 val
	33	289 cys TG:	tyr TAT	Jeu CTT	970 CCT	val GTT	1ys AAA	val GTA	6 000	leu CTT
	a 1 a c c c c c c c c c c c c c c c c c	h is CAC	AAC	ACA ACA	1 ys	TTA TTA	437 438 cys cys TCT TCT	878 CCA		818 GCA
10	P CC	ser his TCC CAC	1ys AAA	leu CTG	phe TT	Jeu CTG		thr Acc	tyr val TAC GTT	thr ala ACT GCA
10	asp asp arg ala GAT GAC AGG GCG		316 lys asp val cys lys asm AAG GAT GTT TGC AAA AAC	ser val val leu leu leu ard TCT GTC GTG CTG CTG AGA	2 Y	40A glu tyr lys phe gln asn ala leu leu val GAG TAC AAA TTC CAG AAT GCG CTG TTA GTT	ser lys AGC AAA	glu lys GAG ĀĀĀ	alu thr GAA ACA	520 glu arq ain ile lys lys ain thr ala GAG AGA CAA ATC AAG AAA CAA ACT GCA
	asb GAT	glu lys GAA AAA	asp val GAT GTT	Jeu CTG	asp GAT	asn AAT	ser AGC			arg ain ile iys iys ain Aga caa atc aag aaa caa
	5 T	leu leu CTG TTG	asp GAT	val	phe TTC	ele CAG	val qly Grc GCC	leu his TTG CAT	asp GAT	1ys AAG
	253 cys TGT	leu CTG	1ys AAG	val	val phe GTG TTC	lys phe gln AAA TTC CAG	val GTG	leu TTG	val	11e ATC
15	253 leu glu cys CTT GAA TGT	pro CCT	ser AGT	367 101	370 tyr ala lys val TAT GCC AAA GTG	lys AA	430 ser arg asn leu qiy lys val qly TCA AGA AAC CTA GGA AAA GTG GCC	val GTG	ale GAA	eg C&
		1ys AA		tyr TAC	နှင် ၁၁၁	tyr TAC	414 664	461 cys TGT	leu CTG	AGA AGA
	250 asp leu CAT CTG	280 91u 6AA	310 val CTT	340 pro asp CCT GAT	369 370 pro his qiu cys tyr CCT CAT GAA TGC TAT	\$10 640 640	430 leu CTA	460 461 gln leu cys CAG TTA TGT	a90 ala cct	520 910 GAG
		278 279 ser lys leu lys glu cys cys AGT AAA CTG AAG GAA TGC TGT	ag F	glu tyr ala arg arg his pro GAA TAT GCA AGA AGG CAT CCT	369 cys 160	91y 66A	AAC		phe ser TTT TCA	514 gys thr leu ser glu lys TGC ACA CTT TCT GAG AAG
20	91. 66.	278 278 160	ala asp GCT GAT	his CAT	a lu GAA	leu	ş ş	AAC	phe TT	g) GVG
20	265 266 oys oys his gly TGC TGC CAT GGA	9 to	ala ala GCT GCT	arg AGG	pro his qiu CCT CAT GAA	glu gln leu GAG CAG CTT	ser TCA	ser val val leu asn TCC GTG GTC CTG AAC	pro cys phe CCA TGC TTT	3er TCT
	246 973 700	Tys AG	ele GCT	arg AGA	pro	alu GAG	glu val GAG GTC	val		leu CTT
	245 glu oys GAA TGC	leu CTG	ser leu TCA TTA	ها. 50۸	ala ala ala asp. GCC GCT GCA GAT	phe TTT	glu	val GTG	arg arg AGG CGA	thr
		1ys AAA		tyr TAT	818 GCA	glu leu GAG CTT	val GTA	36T	879 AGG	514 cys TCC
25	hie thr	ser AGT	leu pro TTG CCT	g lu GAA	ala ala GCC GCT		thr leu val ACT CTT GTA	tyr leu TAT CTA	asa AC	510 his ale asp lie cys CAT GCA GAT ATA TGC
		ser TCC	leu TTG	tyr TAT		392 cys 1GT	thr	tyr 1AT	leu val TTG GTG	asp GAT
	240 lys val AAA GTC	270 ser 11e 7CG ATC	300 ala asp GCT GAC	leu 11G	360 361 0ys cys 1GC 1GT	390 gln asn CAA AAT	420 thr pro ACT CCA	450 glu esp GAA GAC	Jeu 116	a1a GCA
				330 phe 111					480 3er TCC	
	th ACC	gln asp CAA GAT	pro CCT	gly met GGC ATG	glu lys GAG AAG	ile lys ATC AAA	ser TCA	448 cys ala TGT GCA	9 to	thr phe ACC TTC
30	5 7		met ATG	91y 660		•	val GTG		thr	_
	8 5 5 7 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5 5 T 5	265 cys glu asn TGT GAA AAT	g g GAG	Jeu 11G	leu CTA	gin asn leu CAG AAT TTA	gla CAA	5 TO CO	477 cys TGC	phe TTC
	thr ACA	ag W	S. P.	phe TTC	thr	gin asn CAG AAT	. 5 70	met ATG	476 1ys cys AAA TCC	th ACA
35	4.1 CTG	265 cys 1GT	giu asn GAA AAT	val GTC	thr		val GTA	ACA	1 ys	95 84
	1eu	tyr lle TAT ATC		asp GAT	916 83	glu pro GAG CCT	1ys 1ys AAG AAA	1 ys	val thr GTC ACC	al a
	ser lys leu val thr asp TCC AAG TTA GTG ACA GAT	261 ala lys tyr lle GCC AAG TAT ATC	glu val GAA GTG	ala lys asp GCA AAG GAT	351 lys thr tyr glu thr thr AAG ACA TAT GAA ACC ACT	381 val glu glu pro GTG GAA GAG CCT	thr lys lys vel pro ACC AAG AAA GTA CCC	glu ala lys arg met GAA GCA AAA AGA ATG	val GTC	501 glu phe asn ala glu thr GAG TTT AAT GCT GAA ACA
		lys AG			thr ACA	glu GA	thr Acc	38	ACA	phe TTT
	231 481 GT	261 ala GCC	291 818 GCC	321 91u GAG	351 1ys AAG	381 val GTG	411 tyr TAC	pro CCT	471 85P GAC	501 glu GAG

Sab S50	5		559 560 cys lys TCC AAG (1790)	ter TAAAG (1883)	TTAA (2002)	
val lys his lys pro GIG AAA CAC AAG CCC. SSP lys glu thr cys l AI AAG GAG ACC IGC l CTACCATGAGAATAAAGAGA	10		558 55 val glu lys cys cy . GTA GAG AAG TGC TG	leu ter TTA TAA CATCACAFTTA	TCTAAAAACATAAATTTCI	
val lys his lys pro GIG AAA CAC AAG CCC. SSP lys glu thr cys l AI AAG GAG ACC IGC l CTACCATGAGAATAAAGAGA	15	•	550 ssp phe ala ala phe SAT TTC GCT GCT TTT	180 In ala ala leu qiy AA GCT GCC TTA GGC	.TGTAAAGCCAACACCCTG	
val lys his lys pro GIG AAA CAC AAG CCC. SSP lys glu thr cys l AI AAG GAG ACC IGC l CTACCATGAGAATAAAGAGA	20		s ala val met asp a A GCT GTT ATG GAT C	s s val ala ala ser g r GTF GC7 GCA AGT C	GTTTTCTTTTCGTTGG	20 AA (2078)
val lys his lys pro GIG AAA CAC AAG CCC. SSP lys glu thr cys l AI AAG GAG ACC IGC l CTACCATGAGAATAAAGAGA	25	•	lys glu gln leu ly AAA GAG CAA CTG AA	glu qly lys lys lei GAG GGT AAA AAA CT	ATCAAAAGCTTATTCATCI	TGGAAAGAATCTAA
Leu val lys his ly CTC GTG AAA CAC AA CSP asp lys glu th CAC CAT AG GAG AG TAGCCTACCATGGGATT	30			phe ala TTT CCC		CCTCAATTAATAAAAA
531 91u 561 561 7ATCT (CATT	35		ssi glu leu val lya hia ly GAG CTC GTG AAA CAC AA	561 ala asp asp lys glu th GCT GAC GAT AAG GAG AC	ter t ATCTCAGCCTACCATGAGAAT.	CATTTGCCTCTTTCTCTGT

- 10. A nucleotide sequence according to any of claims 6 to 9, in essentially pure form.
- 11. A DNA transfer vector comprising a nucleotide sequence as defined in claim 5.
- 5 12. A DNA transfer vector according to claim 11, transferred to and replicated in a micro-organism.
 - 13. A DNA transfer vector according to claim 12, which is a plasmid.
- 14. A DNA transfer vector according to claim 13,10 wherein the plasmid is pBR322 or YEp6.
 - 15. A process for preparing human serum albumin, which comprises culturing a micro-organism according to claim 5.
- 16. A DNA transfer vector according to any of 15 claims 12 to 14, or a process according to claim 15, wherein the micro-organism is a bacterium or yeast.
 - 17. A vector or process according to claim 16, wherein the bacterium or yeast is <u>E. coli</u> or <u>Saccharomyces</u> <u>cerevisiae</u>.

1/1 = (8) 1 (8) 2.0 586 form TAA Sst | 531/2 Restriction Endonuclease Map of Human Serum Albumin cDNA Clones 493/4 Tag 1 Hinf 1 479/0 419/0 450/1 Hinc II Mbo II **pHA36** 382 Mba 12 Kilobases 325/6 Mbo 11 Q. Hint 1 269/0 Pat 1 (3611) Hpa II (3548) 182/3 Tag ! œ. 4 Hpa II (3548) pHA206 57 Hinf I Pst 1 Mbo 11 Mbo 11 (3611) 16/7 31